Piezoelectric hairpin DNA gas sensors: In silico prediction and experimental proofs

Marcello Mascini [ab], Sara Gaggiotti [a], Flavio Della Pelle [a], Joseph Wang [b], José M. Pingarrón [c] and Dario Compagnone [a]**

[a] *Faculty of Bioscience and Technology for Food, Agriculture and Environment, University of Teramo, 64100, Teramo, Italy*

[b] *Department of Nanoengineering, University of California, San Diego, La Jolla, California 92093, United States*

[c] *Department of Analytical Chemistry, Faculty of Chemistry, University Complutense of Madrid, Ciudad Universitaria s/n, 28040 Madrid, Spain*

Corresponding authors: mmascini@unite.it; dcompagnone@unite.it

Abstract

Hairpin DNA (hpDNA) loops, virtually screened, was used for the first time as artificial trap in solid–gas analysis. The hpDNA loops having unpaired bases were analyzed in silico for the binding to four chemical classes (alcohols, aldehydes, esters and ketones) of volatile organic compounds (VOCs). Two tetramer and pentamer and three hexamer DNA were selected maximizing the recognition properties of the DNA motif between chemical classes. The virtual binding score trend was correlated to the oligonucleotide size increasing of about 25% from tetramer to hexamer. All oligonucleotides showed common trends with best binding scores for alcohols followed by esters, aldehydes and ketones. The seven ssDNA loops (CCAG, TTCT, CCCGA, TAAGT, ATAATC, CATGTC and CTGCAA) were extended with the same double helix stem of four bases (GAAG to 5’ end and CTTC to 3’ end) and covalently bound to gold nanoparticles (AuNPs) using a thiol spacer attached to 5’ end of the hpDNA. HpDNA-AuNP was deposited onto 20 MHz quartz crystal microbalance (QCMs) to realize the gas piezoelectric sensor. An

estimation of relative binding affinities was obtained using different amounts of eight VOCs (ethanol, 3-methylbutan-1-ol, 1-pentanol, octanal, nonanal, ethyl acetate, ethyl octanoate, and butane-2,3-dione) representative of the four chemical classes. Larger DNA loop improved in two orders of magnitude the binding affinity highlighting the DNA size key role. Analyzing data by principal components analysis (PCA), demonstrated the possibility of discriminating VOCs on the basis of molecular weight and functional groups, in agreement with the predicted simulation.

Keywords: Hairpin DNA; In silico screening; Gold nanoparticles; piezoelectric gas sensors, multivariate analysis, VOCs

1 Introduction

In the past decade, DNA was extensively used in sensors design, fabrication, characterization, and application providing new impulses to analytical research (Bettazzi et al. 2017; Rasheed and Sandhyarani 2017). Through the selection of the DNA sequence, a wide variety of analytical applications was proposed, the majority of which were applied to liquid samples. Currently, gas sensors are addressed on mimicking the olfactory system by olfaction-inspired biomaterials and very few propose DNA as functional material (Wasilewski et al. 2017). The first attempts to use DNA in gas sensing started few years ago, particularly by decoration of carbon nanotubes to enhance the affinity and selectivity to gas target analytes (Khamis et al. 2012; Kybert et al. 2013; Su et al. 2013). In a recent work DNA extracted from fish sperm was introduced between a gate dielectric and the organic semiconducting layer to build up an organic field-effect transistor sensor for NO_2 detection (Shi et al. 2016).

To date no gas sensor works had explored the use of hpDNA, neither rationally designed by molecular modelling, for sensing of VOCs. In fact, hpDNA has been used for sensor applications only in liquid media, mainly along with electrochemical transducers (Martín-Fernández et al. 2015; Wang et al. 2014).

In this work, hpDNA conjugated with AuNPs was used as a molecular trap in piezoelectric gas detection. Piezoelectric transducers can monitor the frequency change

of functionalized QCM when gaseous molecules are adsorbed, providing the relationship between mass and resonant frequency shift (Skládal 2016). In gas piezoelectric sensors, the use of AuNP as platform for VOCs binding was found to increase the sensitivity by two orders of magnitude versus monolayer modified QCM (Compagnone et al. 2013).

The new hpDNA-AuNP piezoelectric gas detection strategy, described in this paper, is based on in silico calculation of the hpDNA loop having unpaired bases (Figure 1). In silico rationally designed molecular traps have been demonstrated to possess a strong impact on the development of analytical techniques by minimizing experimental issues such as reagent stability, nonspecific recognition, and separation procedures (Baggiani et al. 2013; Mascini et al. 2013; Narcisi et al. 2011; Uzun and Turner 2016).

Gas sensing computational approach was recently used to reduce the large number of attempts necessary to select the right combination of tools for any given sensing application (Gustafson and Wilmer 2017; Mascini et al. 2017; Pizzoni et al. 2014).

Here, the realization of functional traps was driven by in silico molecular modeling data, with the aim to optimize trial and error analytical protocols by the introduction of prognostic models. We improved the system sensitivity, taking advantage of the AuNPs as the immobilization platform for the hpDNA sequences. The relative binding affinities of the hpDNA loops against different VOCs belonging to relevant chemical classes were evaluated, finding that DNA loop size played a very important role showing an improvement of the binding affinities as the size increased. The pattern recognition of these new sensors was estimated by using the unsupervised multivariate algorithm PCA, a very convenient tool very often used in sensors post-processing analysis (Akamatsu et al. 2017; Compagnone et al. 2015; Imamura et al. 2017). Data obtained demonstrated that the hpDNA sensors were able to discriminate aldehydes and ketones from esters and alcohols, but not between esters and alcohols. The VOCs belonging to both esters and alcohols could be clearly distinguished only through molecular weight.

FIGURE 1

2 Materials and Methods

2.1 In silico screening

The in silico screening procedure was aimed to test the virtual binding affinities of all possible combinations of tetramer, pentamer and hexamer single strain DNA (ssDNA) of the hairpin loop against four chemical classes represented by 50 different compounds. These were: 14 Alcohols ((1S,2R,5R)-2-isopropyl-5-methyl-cyclohexanol; (2S)-propane-1,2-diol; (2Z)-3,7-dimethylocta-2,6-dien-1-ol; (3R)-3,7-dimethylocta-1,6-dien-3-ol; (3R,6Z)-3,7,11-trimethyldodeca-1,6,10-trien-3-ol; (3S)-3,7-dimethyloct-6-en-1-ol; (3S)-oct-1-en-3-ol; (4S,4aR,8aR)-4,8a-dimethyldecalin-4a-ol; 2-Propanol; 3-methylbutan-1-ol; Ethanol; Hex-3-en-1-ol; hexan-1-ol; Terpinen-4-ol); 13 aldehydes ((2E,6Z)-nona-2,6-dienal; (2S)-2-methylbutanal; (2Z)-3,7-dimethylocta-2,6-dienal; (3S)-3,7-dimethyloct-6-enal; (E)-non-2-enal; (Z)-hex-3-enal; 2-methylpropanal; 3-methylbutanal; acetaldehyde; decanal; hexanal;nonanal; octanal); 18 esters (5-methylhexanoate; butanoate; ethyl (2S)-2-methylbutanoate; ethyl 3-methylbutanoate; ethyl acetate; ethyl butanoate; ethyl hexanoate; ethyl octanoate; ethyl propanoate; hexyl acetate; Isopentyl acetate; methyl acetate; methyl butanoate; methyl formate; methyl propanoate; octyl acetate; pentyl butanoate; pentyl pentanoate); 5 ketones (5 molecules: (1S,4S)-1,7,7-trimethylnorbornan-2-one; acetone; butane-2,3-dione; cyclopentadecanone; pentane-2,3-dione). The 50 molecules were selected in order to have different functional groups and dimensions.

The entire DNA library of tetramer, pentamer and hexamer ssDNA was generated using Hyperchem 8.0.5 software on a Microsoft Windows 10 laptop. Calculations of the in silico screening process, including molecular docking run and data preparation were performed using a desktop computer with 19 processors Intel Xeon X5690 at 3.47 GHz each, with 94.5 GiB RAM, running Kernel Linux 2.6.32-642.1.1el6.x86_64, GNOME 2.28.2. Tools from OpenEye Scientific Software package under academic license, were used at different stages of the in silico procedure. VOCs were obtained via LEXICHEM 2.1.0 package, by converting ligands standard IUPAC names into their corresponding structures (LEXICHEM version 2.1.0). SZYBKI 1.5.7 with default parameterization was used to optimize molecular geometries (SZYBKI version 1.5.7). Conformational space

for both ssDNA and VOCs was taken into account with OMEGA 2.4.6 (Hawkins and Nicholls 2012; Hawkins et al. 2010; OMEGA version 2.4.6). Multi-conformer rigid body docking was carried out using OEDocking 3.0.0, having Chemgauss4 as scoring function (Kelley et al. 2015; OEDocking version 3.0.0). Structures visualization and generation of molecular surfaces were performed using VIDA 4.2.1(VIDA version 4.1.1).

The entire DNA molecular surface was included in the active site box defining the area where VOCs were expected to bind. For each ssDNA receptor, a dedicated box (10– 20 nm^3) was generated. The time elapsed for processing each DNA conformer was about 2 min per processor, from the initial 3D structures generation to final docking results. Ten conformers per ssDNA and a maximum of 200 conformers for each of the 50 VOCs were considered. The binding score average for each DNA was calculated over all the conformers. The entire process was automated using a bash script and using a freeware BASIC-like scripting language (AutoIT V3) for post processing data analysis.

2.2 Experimental procedure

All the reagents and the eight VOCs were purchased from Sigma-Aldrich (Italy). The eight VOCs (ethanol, 3-methylbutan-1-ol, 1-pentanol, octanal, nonanal, ethyl acetate, ethyl octanoate, and butane-2,3-dione) were of analytical grade. HpDNA having unpaired tetramer loop were purchased from Thermo Fischer Scientifics (Italy); the others were from Integrated DNA technologies (USA). Standard desalting purification oligonucleotides were bought with a thiol spacer having six carbons.

The piezoelectric measurements were carried out using an Enose-UTV from Sensor group, University of Rome Tor Vergata (Italy). 20 MHz QCM sensors, were from KVG GmbH (Germany).

Colloidal AuNPs were synthesized using the trisodium citrate reduction method (Frens 1973).In brief, 50 mL of 0.3 mM tetrachloroauric acid solution was stirred vigorously and heated. When the solution was boiling, 1.5 mL of 40 mM trisodium citrate solution was added. After boiling for 20 min, the color of the solution turned from clear liquid to wine red. Then, the solution was cooled down to 4 °C for future use. Ultraviolet-visible

spectrophotometry was performed to confirm the AuNPs formation and verify the AuNPs dispersion.

Immobilization of the oligonucleotides on the AuNP surface was carried out covalently using a C6 thiol modifier group attached to 5' phosphate end of the hpDNA. Each hpDNA was dissolved in deionized water and added to 1mL of the AuNPs colloidal solution at a final concentration of 27.1 μM. The hpDNA-AuNPs colloidal suspensions were incubated at +5 °C for 12 hours. HpDNA-AuNPs were then centrifuged at 13000 rpm for 30 min at 4 °C. The colorless supernatant was discarded and the solid pellet was resuspended in 1 mL of deionized water. All steps were monitored via UV-Vis spectrophotometry.

The QCM sensors modification was achieved by drop casting 5 μL of the hpDNA-AuNP - suspension on each side of the crystal and let dry for few minutes. Before the first use, the QCM sensors were completely dried under N_2 at a flow rate of 2L/h and stored at room temperature in the dark when not in use.

The piezoelectric measurements were carried out using N_2 as carrier gas at a flow rate of 2 L/h. The frequency shift (ΔF), taken as analytical signal, was recorded after adding different volumes of the VOCs to a 100 mL laboratory glass container, connected to a tubing system with three-way stop-cocks. Before measuring, the three-way stop-cocks were closed for 10 min allowing the solvent to evaporate completely at 45 °C and then to equilibrate the head-space at 25 °C in stationary conditions. The stop-cocks were then opened to carry the head-space of the 100 mL glass bottle to the measuring sensor chamber. Steady state was reached between 100 and 200 s after opening the stop-cocks. After each measurement, a complete recovery of the signal was achieved under N_2 flow in about 400 s. The piezoelectric sensorgram was similar for all hpDNA-AuNP and VOCs, showing a rapid decrease of the signal after the stop-cocks opening, followed by a slower raise up to the steady state. The ΔF, was recorded for all VOCs before desorption.

Piezoelectric responses dataset was analyzed by the unsupervised multivariate technique principal component analysis (PCA) using MatLab R2011 (USA). Dataset were autoscaled (zero mean and unitary variance) before analysis. PCA was applied to

inspect the multivariate data structure by decomposing a data matrix of eight rows (the VOCs) and seven columns (the hpDNA-AuNP sensors).

3 Results and discussion

3.1 –In silico screening: ssDNA vs chemical classes

The binding properties of the ssDNA library were calculated against 50 VOCs molecules belonging to four different chemical classes (alcohols, aldehydes, esters and ketones). All possible combinations of tetramer, pentamer and hexamer DNA were tested. The ssDNA library was built using in every position of the sequence the natural bases adenine (A), cytosine (C), guanine (G) and thymine (T) resulting in a total of 5376 library elements. DNA was preferred over RNA because this is less stable due to base-catalyzed hydrolysis and auto-hydrolysis in single strand structure.

The molecular docking functions used in this work screened out compounds that potentially interacted with the binding site predominantly through non-covalent interactions, particularly hydrogen bonds. Therefore, only the hpDNA loops having unpaired bases were virtually screened. Table 1 reports the statistical summary of the binding scores calculated for the three groups tetramer, pentamer and hexamer DNA towards the four chemical classes. The binding score was reported as the average calculated over 10 conformers for each DNA. The score values were calculated using chemgauss4 scoring function, thus lower values represented higher ssDNA–ligand affinity. The oligonucleotides virtual binding score trend was correlated to the oligonucleotide size for all chemical classes, with values increasing of about 25% from tetramer to hexamer. All oligonucleotides showed common trends with best binding scores for alcohols followed by esters, aldehydes and ketones. These later compounds showed a decrease in binding score of 50% compared to alcohols. The minimum-maximum dynamic range for each chemical class was quite narrow for tetramers becoming relevant only for hexamers DNA. In all cases, average and median were very close to each other demonstrating a good symmetry in normal distribution.

TABLE 1

The structural analysis of the four bases in the 5% top ranked structures of the tetramer, pentamer and hexamer unpaired DNA versus the four chemical classes are reported in Table 2. The structural data exhibited a very high level of similarity in DNA bases distribution. Top ranked tetramer and pentamer DNA had higher amount of adenine and thymine, however, in the hexamer DNA the occurrence of both purines was higher than pyrimidines.

Due to the small combinations generated by only four different DNA bases, the binding difference within the DNA library was likely due to the steric/conformational effects. Increasing the DNA size enhanced the internal flexibility of specific DNA regions and the target accessibility to the binding box conformational space.

Figure 2 shows the specific positions of the DNA bases cooperatively improved the target binding. Top binding scores were obtained when DNA docked VOCs with a saddle shaped binding pocket, allowing oligonucleotide to bury the entire ligand in its surface. On the other hand, a docking with a planar interaction surface was not efficient confirming that the amount of freedom degree of the single bases to move around the DNA backbone was the major effect to explain the binding score data; this is particularly true for hexamer DNA where the probability of synergic cooperation is higher.

The results of the virtual screening were used to select some oligonucleotides with different affinities for the VOCs in order to evaluate their potential applicability in gas analysis by using QCM sensors. The selection was finalized to maximize the recognition properties of DNA motif between chemical classes. Two tetramer and pentamer and three hexamer DNA were chosen to be used as loops of the hpDNA. Table 3 reports the binding score of the DNA versus the VOCs selected in experimental part. To emphasise the differences between chemical class average and single compounds of the same class, the binding score average obtained by the simulations of the ssDNA versus the chemical classes (14 alcohols, 13 aldehydes, 18 esters and 5 ketones) is also reported.

TABLE 2

FIGURE 2

TABLE 3

The selected oligonucleotides have the same trend of the entire DNA library with better interaction for alcohols followed by esters and aldehydes, and ketones with lowest interactions. According to the binding score data, all the DNA sequences exhibited similar trend for alcohols except for ethanol; binding scores varied significantly for the interaction with esters and aldehydes and, in the case of some hexamer, also for ketones.

The oligonucleotide TTCT showed a good interaction particularly for aldehydes and ethyl octanoate. Moreover, the other tetramer CCAG exhibited a clear difference between alcohols and the other VOCs selected in experimental part. The pentamer TAAGT was selected because of the very low interaction with all the molecules compared to its counterpart CCCGA which showed in every case almost two-fold more interaction energy.

A clear difference in affinity scores was observed using the hexamer DNA. As reported also considering the entire DNA library, increasing the number of bases, there was a considerable increase of docking scores. The hexamer ATAATC showed better binding score than the other oligonucleotide receptors for most of the ligands and, in particular, for ethyl octanoate and both aldehydes. This hexamer and CATGTC exhibited the same pattern in docking the alcohols, aldehydes and esters, showing a significant difference between the small molecules, ethanol and ethyl acetate, and the other molecules. All oligonucleotides exhibited affinity properties inversely correlated to molecular weight except the third hexamer DNA CTGCAA with good affinity for ethanol. The latter hexamer had, for all VOCs, half interaction energy compared to the other two hexamer DNA.

It should be noted that the same stem DNA sequence was used for the realization of the hpDNAs in order to evaluate the contribution of the loop. Thus, some oligonucleotides, particularly in hexamer DNA, were discarded due to stem-loop intramolecular base pairing.

3.2 AuNPs-DNA functionalization and QCM sensors modification

After in silico screening each of the seven ssDNA was extended attaching to the 5' end the sequence GAAG and to the 3' end the sequence CTTC. This allowed folding to give hpDNA. Each secondary structure was analyzed using the Mfold Web Server (www.unafold.rna.albany.edu) to check the stem-loop intramolecular base pairing. All selected DNA had unpaired loop in standard condition.

The AuNP functionalization with hpDNA was monitored by UV-Vis spectroscopy. The amount of hpDNA for the AuNPs functionalization was selected testing different concentrations (0.136, 0.271, 0.678 and 1.355 μM.) of hpDNA.

Figure 3A displays the UV-VIS spectra after AuNP functionalization with hpDNA (TTCT as loop) before centrifugation. Similar absorption spectra were obtained in the 350 to 800 nm range for bare AuNPs and all the different amounts of hpDNA-AuNPs demonstrating that the functionalization did not cause AuNPs aggregation. Similar results were obtained using the other DNA loops.

Figure 3B and 3C show the UV-Vis spectra of the resuspended hpDNA-AuNPs and the relative supernatant after centrifugation, respectively. Centrifugation was essential to remove chemicals excess after nanoparticles functionalization and often leads to a loss of nanomaterial and causes aggregation of nanoparticles.

The hpDNA-AuNPs resuspended in water showed similar spectra, with the same absorption maximum, indicating high stability of the suspensions due to the interactions among hpDNA-AuNPs and indirectly surface modification of the AuNPs. Unmodified AuNPs were not easily resuspended in water (black line in Figure 3B), indicating an irreversible aggregation due to the centrifugation step. The spectra showed clearly that the AuNPs were stabilized by the negatively charged DNA that acted an electrostatic repulsing capping agent among the AuNPs as clearly reported in the literature (Baldock and Hutchison 2016; Xu et al. 2016). The presence of DNA onto the nanoparticles was confirmed by the presence of a clear peak at 260 nm.

The spectra of the supernatant (Figure 3C) show the difficulty to clean-up classical citrate capped AuNPs with the presence of an absorption peak at 520 nm. The peak due to AuNPs for the supernatant of the hpDNA-AuNPs was significantly lower and a

peak at 260 nm increasing with the hpDNA concentration used in AuNP functionalization was also observed. A final concentration of 0.675 μM DNA was selected to implement the gas sensors.

FIGURE 3

The 20 MHz QCM sensors surfaces were modified by drop-casting of 2.5 μL of hpDNA-AuNP suspension on each side of the crystal and let drying at room temperature. This procedure was repeated to assess the maximum loadable amount. Every 2.5 μL addition of hpDNA-AuNP suspension on each side of the sensor led to a variation of approximately 2.5 KHz. After four times (20 μL total volume) QMC crystals frequency crashed and no variation was detectable. Thus, a total amount of 20 μL of hpDNA-AuNP suspension was selected for further work, leading to a variation of 10 KHz in all cases.

3.3 QCM sensors response to VOCs

QCMs frequency shifts (ΔF) were used to calculate the relative experimental binding constants of the eight VOCs and to assess the correlation between the virtual screening and real binding data. For this reason, pure VOCs were tested by using N_2 as carrier gas directly in the measuring chamber.

The relative binding affinities of the complex hpDNA VOC were calculated by adding to the sample glass container different amount of pure solvent. After solvent evaporation, the real-time binding to the surface was estimated by recording the frequency shift. The latter was used for a quantitative evaluation of the mass captured by the QCM sensors using the oscillation constant (Kq=−4.8 Hz/ng), to calculate, consequently, the moles bound by the sensor. Figure 4 shows, as an example, the frequency shifts measured with the sensor modified with CTGCAA as loop, for different amounts of 1-pentanol. The thermodynamic equilibrium was reached between 100 and 200 sec after the start of the measurement. The adsorption kinetics was similar for all the VOCs tested; the hpDNA-AuNP-QCMs showed a rapid decrease of the signal after the stop-cocks opening, followed by a slower raise up to the thermodynamic equilibrium adsorption.

The frequency shift (ΔF), taken as analytical signal, was recorded for all cases before desorption.

The bound compound was determined by considering 1:1 complexation stoichiometry. Using Scatchard model, the ratio between bound and free compound versus the bound was plotted and the relative binding affinity was calculated by linear regression fitting. The results are reported in Table 4.

FIGURE 4

Despite their different structure both tetramer DNA loops had very similar binding affinity for all VOCs. TTCT exhibited slight better affinity for aldehydes leading to a significant correlation with simulated results. On the contrary, there was no correlation between the other DNA tetramer loop and virtual data for the low and high experimental affinities obtained for 1-pentanol and ethyl octanoate, respectively. These two molecules and 3-methylbutan-1-ol were bound by the pentamer CCCGA with affinity one order of magnitude higher than both DNA tetramers. The other DNA pentamer loop, TAAGT, had the lowest binding affinity for all molecules. The correlation coefficient of this pentamer DNA was only 0.37 because of alcohols results, slightly high in silico but low in experimental.

The DNA loop size played an important role in the observed experimental behavior improving the binding affinities, as revealed by the DNA hexamer loop binding data. Both hpDNA having as loop ATAATC and CATGTC showed a significant interaction with ligands, which was approximately two fold higher than the smaller DNA loop. This was in good agreement with the prediction by virtual screening. Strong interaction with larger molecular weight molecules such as 1-pentanol, octanal, nonanal and ethyl octanoate was observed. Virtual screening partially predicted the completely different behavior of the other DNA hexamer, CTGCAA, considering the relative good affinity for alcohols but not the smaller affinity for the other molecules compared to the other DNA hexamers. The different responses of these DNA hexamers to the VOCs emphasized the importance of the chemical nature of the DNA loop chemical nature. Such

heterogeneous data set demonstrated that the binding affinities did not depend on the presence of the stem that was the same for all hpDNA.

TABLE 4

The inter-relationships between the sequence-specific responses of hpDNA to VOCs were highlighted considering all sensors measurements together in multivariate analysis format. Data structure was analyzed by means of PCA. The data set was represented by the hpDNA-AuNPs-QCMs frequency shifts obtained using 900 μmoles of each VOC.

Before applying PCA the data were autoscaled and then analyzed making use of unsupervised PCA. Figure 3 shows the biplot of the first two principal components. The first component represented 46.84% of the variance, the second 28.19% displaying together a cumulative variance of 75.03%.

The score points in blue representing the new coordinates of the VOCs were interpreted assuming that close distance in plot plane is a measure of the similitude between samples. PC 1 well separated both aldehydes and ketone butane-2,3-dione from alcohols and esters. PC 2 highlighted the differences within alcohol and ester classes grouping the low molecular weight molecules ethanol and ethyl acetate. The PC 2 was the most influenced by the loadings (in red) representing the contribution of each DNA sensor to the principal components. This axis highlighted the differences among sensors. Both pentamer DNA contributed significantly to the separation of the small alcohol and ester to the other family members. On the other hand, the hexamer ATAATC and the tetramer TTCT played an important role in clustering on the PC 2 the molecules with higher molecular weight confirming the predictions obtained by virtual data.

It is important to note that hpDNA-AuNPs-QCM sensors could not clearly discriminate esters and alcohols, but they could clearly separate molecules based on the molecular weight. The other two hexames had very similar pattern recognition performance contributing only in spreading the VOCs on PC 1. The PCA algorithm highlighted that the DNA sensors can be effectively applied to those cases where the difference between VOC patterns plays a crucial role in classification purposes.

FIGURE 5

4 Conclusions

This work contributes to the growth of the DNA applications in biotechnological and analytical field. For the first time, the interaction between hpDNA loops and VOCs were rationally calculated by virtual assessment and then experimentally tested. A good matching between in silico selection and experimental results was found with almost all hpDNA tested.

Multivariate data elaboration showed that beyond interesting differences between aldehydes and ketones, esters and alcohols could be clearly discriminated based only on the molecular weight.

The key parameter for increasing the affinities of sensors versus VOCs was found to be the size of the DNA loop within the hairpin structure.

This work represents the starting point for hpDNA selection used as artificial traps in solid-gas sensors. In near future, taking advantage of the fast progress in computing, much bigger ssDNA loops with more complicated shapes could be screened in short times, tailoring the efficiency and effectiveness of the gas analysis.

5 Acknowledgement

This work was supported in part from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 656687. The authors would like also to acknowledge NBCR award number (P41 GM103426).

6 References

Akamatsu, T.I.T., Tsuruta, A., Shin, W., 2017. Selective Detection of Target Volatile Organic Compounds in Contaminated Humid Air Using a Sensor Array with Principal Component Analysis. Sensors 17(7), 1662.

Baggiani, C., Biagioli, F., Anfossi, L., Giovannoli, C., Passini, C., Giraudi, G., 2013. Effect of the mimic structure on the molecular recognition properties of molecularly imprinted polymers for

ochratoxin A prepared by a fragmental approach. Reactive and Functional Polymers 73(6), 833-837.
Baldock, B.L., Hutchison, J.E., 2016. UV–Visible Spectroscopy-Based Quantification of Unlabeled DNA Bound to Gold Nanoparticles. Analytical chemistry 88(24), 12072-12080.
Bettazzi, F., Marrazza, G., Minunni, M., Palchetti, I., Scarano, S., 2017. Biosensors and Related Bioanalytical Tools. Comprehensive Analytical Chemistry.
Compagnone, D., Faieta, M., Pizzoni, D., Di Natale, C., Paolesse, R., Van Caelenberg, T., Beheydt, B., Pittia, P., 2015. Quartz crystal microbalance gas sensor arrays for the quality control of chocolate. Sensors and Actuators B: Chemical 207, 1114-1120.
Compagnone, D., Fusella, G.C., Del Carlo, M., Pittia, P., Martinelli, E., Tortora, L., Paolesse, R., Di Natale, C., 2013. Gold nanoparticles-peptide based gas sensor arrays for the detection of food aromas. Biosensors & bioelectronics 42, 618-625.
Frens, G., 1973. Controlled nucleation for the regulation of the particle size in monodisperse gold suspensions. Nature 241(105), 20-22.
Gustafson, J.A., Wilmer, C.E., 2017. Computational Design of Metal–Organic Framework Arrays for Gas Sensing: Influence of Array Size and Composition on Sensor Performance. The Journal of Physical Chemistry C 121(11), 6033-6038.
Hawkins, P.C., Nicholls, A., 2012. Conformer generation with OMEGA: learning from the data set and the analysis of failures. Journal of chemical information and modeling 52(11), 2919-2936.
Hawkins, P.C., Skillman, A.G., Warren, G.L., Ellingson, B.A., Stahl, M.T., 2010. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. Journal of chemical information and modeling 50(4), 572-584.
Imamura, R., Murata, N., Shimanouchi, T., Yamashita, K., Fukuzawa, M., Noda, M., 2017. A Label-Free Fluorescent Array Sensor Utilizing Liposome Encapsulating Calcein for Discriminating Target Proteins by Principal Component Analysis. Sensors 17(7), 1630.
Kelley, B.P., Brown, S.P., Warren, G.L., Muchmore, S.W., 2015. POSIT: flexible shape-guided docking for pose prediction. Journal of chemical information and modeling 55(8), 1771-1780.
Khamis, S., Jones, R., Johnson, A.C., Preti, G., Kwak, J., Gelperin, A., 2012. DNA-decorated carbon nanotube-based FETs as ultrasensitive chemical sensors: Discrimination of homologues, structural isomers, and optical isomers. Aip Advances 2(2), 022110.
Kybert, N.J., Lerner, M.B., Yodh, J.S., Preti, G., Johnson, A.C., 2013. Differentiation of complex vapor mixtures using versatile DNA–carbon nanotube chemical sensor arrays. ACS nano 7(3), 2800-2807.
LEXICHEM, version 2.1.0. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.
Martín-Fernández, B., de-los-Santos-Álvarez, N., Lobo-Castañón, M.J., López-Ruiz, B., 2015. Hairpin-based DNA electrochemical sensor for selective detection of a repetitive and structured target codifying a gliadin fragment. Analytical and bioanalytical chemistry 407(12), 3481-3488.
Mascini, M., Montesano, C., Sergi, M., Perez, G., De Cicco, M., Curini, R., Compagnone, D., 2013. Peptides trapping cocaine: docking simulation and experimental screening by solid phase extraction followed by liquid chromatography mass spectrometry in plasma samples. Analytica chimica acta 772, 40-46.
Mascini, M., Pizzoni, D., Perez, G., Chiarappa, E., Di Natale, C., Pittia, P., Compagnone, D., 2017. Tailoring gas sensor arrays via the design of short peptides sequences as binding elements. Biosensors & bioelectronics 93, 161-169.
Narcisi, V., Mascini, M., Perez, G., Del Carlo, M., Tiscar, P.G., Yamanaka, H., Compagnone, D., 2011. Electrochemical genosensors for the detection of Bonamia parasite. Selection of single strand-DNA (ssDNA) probes by simulation of the secondary structure folding. Talanta 85, 1927–1932.

OEDocking, version 3.0.0. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.
OMEGA, version 2.4.6. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.
Pizzoni, D., Mascini, M., Lanzone, V., Del Carlo, M., Di Natale, C., Compagnone, D., 2014. Selection of peptide ligands for piezoelectric peptide based gas sensors arrays using a virtual screening approach. Biosensors & bioelectronics 52, 247-254.
Rasheed, P.A., Sandhyarani, N., 2017. Carbon nanostructures as immobilization platform for DNA: A review on current progress in electrochemical DNA sensors. Biosensors & bioelectronics 97, 226.
Shi, W., Yu, X., Zheng, Y., Yu, J., 2016. DNA based chemical sensor for the detection of nitrogen dioxide enabled by organic field-effect transistor. Sensors and Actuators B: Chemical.
Skládal, P., 2016. Piezoelectric biosensors. TrAC Trends in Analytical Chemistry 79, 127-133.
Su, H.C., Zhang, M., Bosze, W., Lim, J.H., Myung, N.V., 2013. Metal nanoparticles and DNA co-functionalized single-walled carbon nanotube gas sensors. Nanotechnology 24(50), 505502.
SZYBKI, version 1.5.7. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.
Uzun, L., Turner, A.P., 2016. Molecularly-imprinted polymer sensors: Realising their potential. Biosensors and Bioelectronics 76, 131-144.
VIDA, version 4.1.1. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.
Wang, L., Hua, E., Liang, M., Ma, C., Liu, Z., Sheng, S., Liu, M., Xie, G., Feng, W., 2014. Graphene sheets, polyaniline and AuNPs based DNA sensor for electrochemical determination of BCR/ABL fusion gene with functional hairpin probe. Biosensors and Bioelectronics 51, 201-207.
Wasilewski, T., Gębicki, J., Kamysz, W., 2017. Advances in olfaction-inspired biomaterials applied to bioelectronic noses. Sensors and Actuators B: Chemical.
Xu, Q., Lou, X., Wang, L., Ding, X., Yu, H., Xiao, Y., 2016. Rapid, Surfactant-Free, and Quantitative Functionalization of Gold Nanoparticles with Thiolated DNA under Physiological pH and Its Application in Molecular Beacon-Based Biosensor. ACS applied materials & interfaces 8(40), 27298-27304.

Highlights of: Piezoelectric hairpin DNA gas sensors

HpDNA loops were screened in silico for the binding to four chemical classes.

Seven HpDNA loops were extended with the same double helix stem and covalently bound to AuNPs.

HpDNA-AuNP was deposited onto 20 MHz QCMs to realize gas piezoelectric sensors.

Relative binding affinities were obtained using different amounts of eight VOCs.

DNA size played a key role, larger DNA loop improved the binding affinity.

HpDNA-VOCs inter-relationships were analyzed by multivariate analysis.

Discrimination of aldehydes and ketones from esters and alcohols was obtained.

Esters and alcohols were clearly distinguished only through molecular weight.

Figure(s)

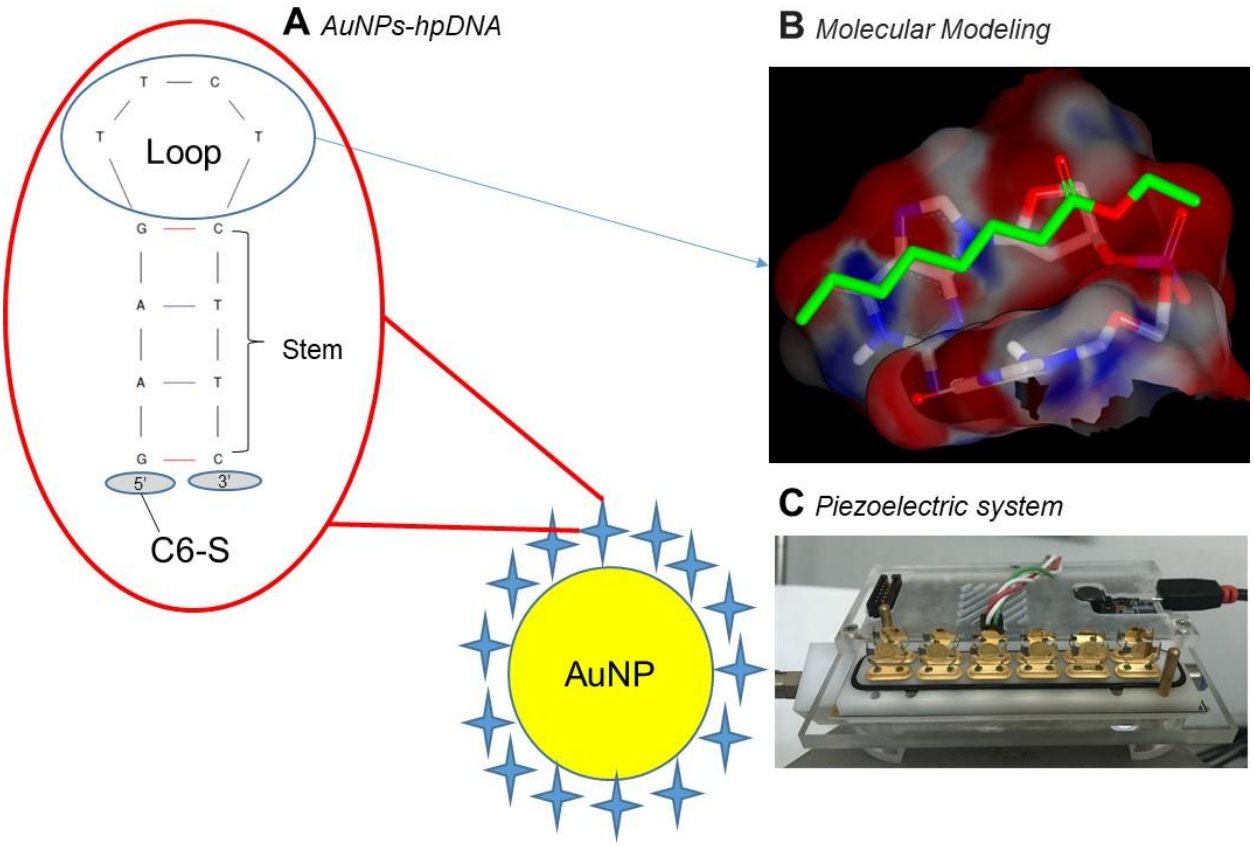

Figure 1. Schematic illustration of A) the hpDNA bound to AuNP via thiol spacer, B) the modelling of the loop used in virtual simulation, C) the piezoelectric quartz crystal microbalance system.

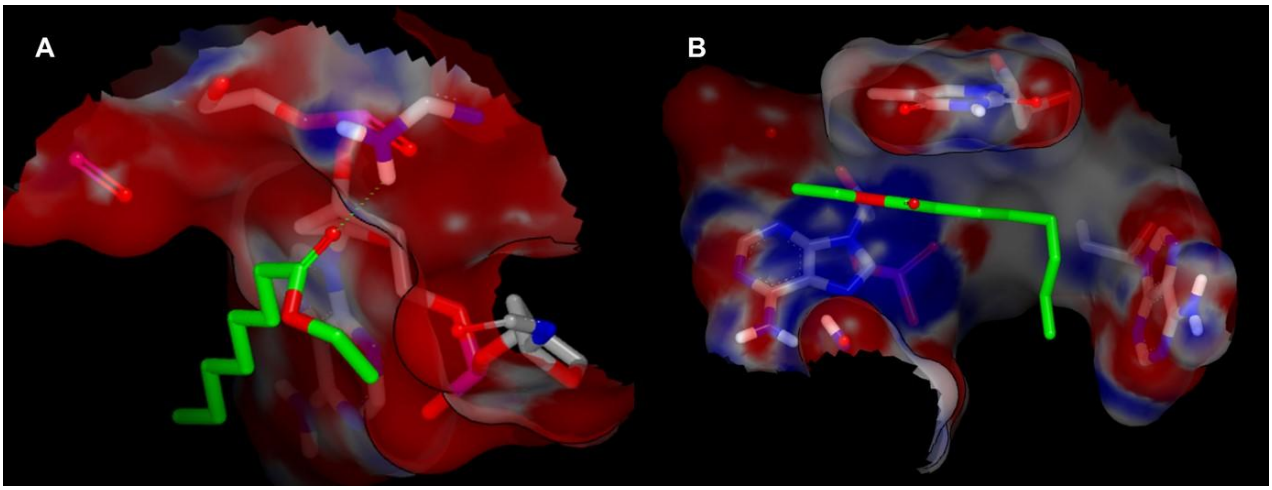

Figure 2. Electrostatic molecular surfaces of the ssDNA CTGCAA, with a planar interaction surface (binding score -2.26 Kcal/mol) (A), and ATAATC with a saddle shaped binding pocket (binding score -6.28 Kcal/mol) (B) in complex with the ethyl octanoate (highlighted in green).

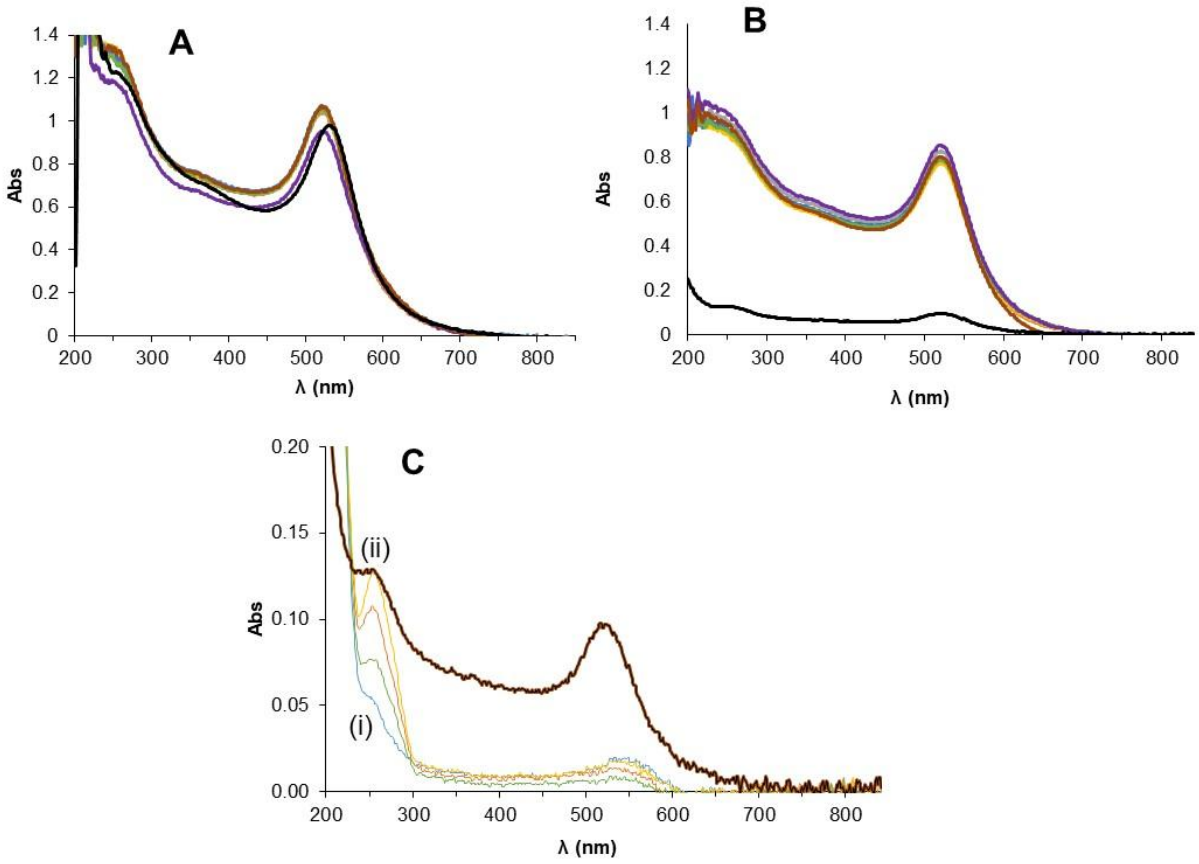

Figure 3. Spectra obtained from the bare AuNPs (in black) and different AuNP- hpDNA loadings before centrifugation step and water resuspension (A). Spectra obtained from the bare AuNPs (in black) and the different hpDNA-AuNP after the centrifugation and water resuspension (B). Spectra for the supernatant obtained after centrifugation from the bare AuNPs (in black) and different concentrations of hpDNA used in AuNP functionalization. From (i) to (ii) the spectra obtained using the four hpDNA concentrations: 0.136, 0.271, 0.678 and 1.355 μM.

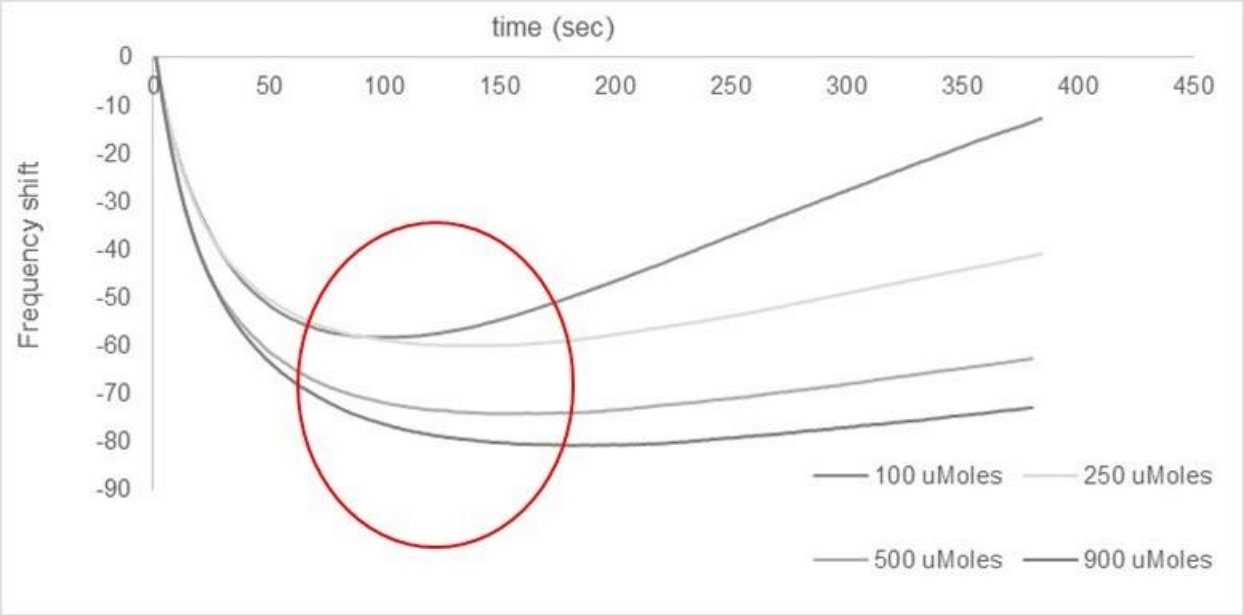

Figure 4. Frequency shifts recorded by the sensor modified with CTGCAA after introducing in the 100-mL glass bottle different μmol amounts of 1-pentanol. In all cases, the relative binding affinity between hpDNA and VOCs were calculated taking the frequency shift before desorption, corresponding to the thermodynamic equilibrium of the compound adsorbed on the QCM surface modified with hpDNA-AuNP.

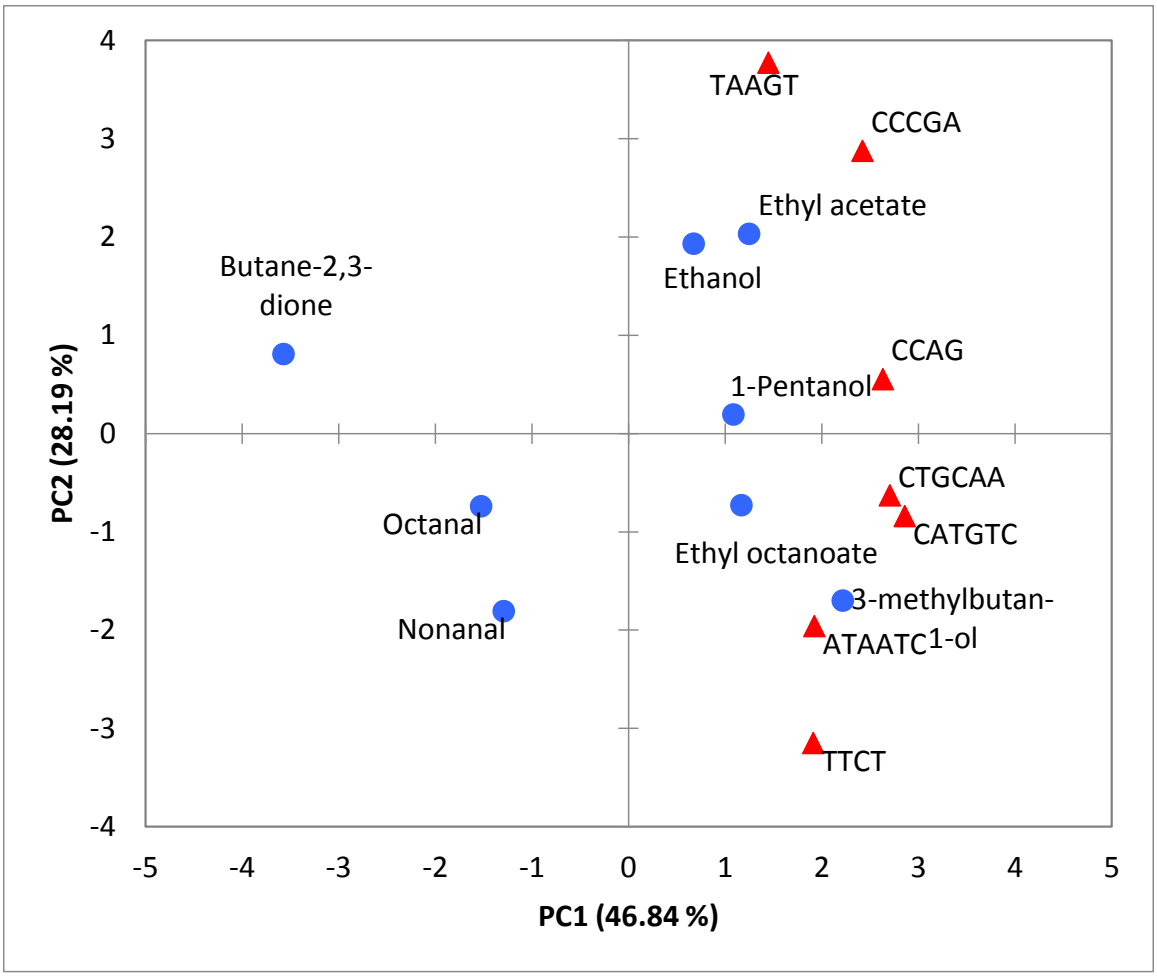

Figure 5. PCA of the piezoelectric responses of hpDNA-AuNP sensors obtained using 900 μmoles of each VOC. The biplot (Score and loading) of the first two principal components showed 75.03% of the cumulative variance. Data were autoscaled before PCA.

Table(s)

	Alcohols	Aldehydes (Kcal/mol)	Esters	Ketones
		Tetramer DNA		
max	-2.32	-1.60	-1.65	-1.17
min	-4.23	-3.48	-3.44	-2.69
av	-3.23	-2.40	-2.41	-1.78
median	-3.21	-2.39	-2.41	-1.77
		Pentamer DNA		
max	-2.39	-1.55	-1.58	-1.24
min	-4.78	-3.89	-4.00	-3.16
average	-3.75	-2.91	-2.92	-2.24
median	-3.74	-2.87	-2.88	-2.21
		Hexamer DNA		
max	-2.47	-1.64	-1.68	-1.22
min	-5.55	-4.69	-4.77	-3.96
average	-4.07	-3.17	-3.19	-2.42
median	-4.12	-3.21	-3.23	-2.45

Table 1. Statistical summary of the binding score average (Kcal/mol), representing the virtual binding energy of tetramer, pentamer and hexamer unpaired hpDNA towards the four chemical classes tested. The binding score average of each DNA was calculated over 10 conformers, the coefficient of variation was in all cases lower than 10%.

	Alcohols	Aldehydes	Esters	Ketones
		(%)		
Tetramer DNA				
A	29	38	38	30
C	21	20	20	20
G	20	18	18	21
T	30	25	25	29
Pentamer DNA				
A	31	31	31	24
C	23	21	21	24
G	17	22	23	22
T	29	26	25	30
Hexamer DNA				
A	31	31	32	28
C	17	19	18	19
G	31	29	31	32
T	20	21	20	22

Table 2 Structural analysis reporting the occurrence percentage of the four bases in the 5% top ranked structures of the tetramer, pentamer and hexamer DNA versus the four chemical classes tested.

	CCAG	TTCT	CCCGA	TAAGT	ATAATC	CATGTC	CTGCAA
Ethanol	-2.72	-2.41	-2.77	-2.28	-3.21	-3.09	-3.24
3-methylbutan-1-ol	-3.23	-2.77	-3.26	-2.74	-4.63	-4.12	-2.72
1-pentanol	-3.30	-3.22	-3.53	-2.48	-5.14	-4.21	-2.82
Alcohols	***-2.94***	***-3.12***	***-3.38***	***-2.39***	***-5.55***	***-4.75***	***-3.01***
Octanal	-2.39	-3.07	-2.62	-1.54	-5.13	-3.71	-2.40
Nonanal	-2.32	-3.24	-2.56	-1.53	-5.54	-4.05	-2.56
Aldehydes	***-2.25***	***-2.78***	***-2.42***	***-1.55***	***-4.67***	***-3.39***	***-2.27***
Ethyl acetate	-2.17	-2.19	-2.49	-1.51	-3.64	-2.68	-1.90
Ethyl octanoate	-1.95	-3.14	-2.49	-1.46	-6.28	-4.20	-2.26
Esters	***-2.19***	***-2.73***	***-2.53***	***-1.58***	***-4.60***	***-3.37***	***-2.26***
Butane-2,3-dione	-2.02	-2.51	-1.88	-1.44	-3.13	-2.17	-1.95
Ketones	***-1.84***	***-2.17***	***-1.94***	***-1.24***	***-3.55***	***-2.43***	***-1.81***

Table 3. Binding score average (Kcal/mol) of the tetramer, pentamer and hexamer DNA versus the VOCs tested. In italic-bold, the binding score average obtained by the simulations of the ssDNA versus the chemical classes (14 alcohols, 13 aldehydes, 18 esters and 5 ketones) was also reported. The conformers and coefficient of variation were the same than in Table 1.

	CCAG			TTCT			CCCGA			TAAGT			ATAATC			CATGTC			CTGCAA		
	Experimental K binding $\times 10^3$ ($Moles^{-1}$)																				
Ethanol	4.3	±	0.6	2.3	±	0.1	7.4	±	0.5	0.9	±	0.1	6.5	±	0.5	8.4	±	0.8	18.4	±	2.8
3-methylbutan-1-ol	6.7	±	0.5	6.5	±	0.6	23.6	±	3.3	7.0	±	0.6	91.2	±	11.9	70.1	±	3.5	21.7	±	2.2
1-pentanol	3.3	±	0.4	3.1	±	0.3	60.3	±	7.8	2.0	±	0.1	530.6	±	58.4	352.8	±	17.6	10.3	±	1.3
Octanal	2.6	±	0.3	6.3	±	0.6	4.7	±	0.3	2.5	±	0.1	582.2	±	81.5	348.0	±	41.8	3.8	±	0.3
Nonanal	2.6	±	0.2	7.9	±	0.6	6.5	±	0.8	1.9	±	0.1	283.1	±	14.2	97.9	±	8.8	4.8	±	0.4
Ethyl acetate	0.9	±	0.1	2.4	±	0.3	9.8	±	1.5	3.1	±	0.4	19.6	±	1.2	17.8	±	2.3	0.9	±	0.0
Ethyl octanoate	8.8	±	0.8	12.8	±	0.6	43.1	±	4.3	2.5	±	0.2	519.2	±	41.5	345.8	±	17.3	6.9	±	0.5
Butane-2,3-dione	1.4	±	0.2	3.4	±	0.3	2.9	±	0.3	3.4	±	0.4	9.2	±	1.1	4.1	±	0.3	2.5	±	0.4
Correlation with Simulated results	0.16			0.63			0.65			0.37			0.83			0.67			0.80		

Table 4. HpDNA-AuNP sensors relative binding affinities vs the VOCs, estimated using piezoelectric response. The correlation coefficient between experimental and simulated binding is reported in the last row. The standard deviation was calculated using three measurements taken in three different days.